DOCUMENT RESUME

ED 208 388 CS 006 343

AUTHOR Snow, David P.; Coots, James H.
TITLE Sentence Perception in Listening and Reading. Technical Note.
INSTITUTION Southwest Regional Laboratory for Educational Research and Development, Los Alamitos, Calif.
SPONS AGENCY National Inst. of Education (ED), Washington, D.C.
REPORT NO SWRL-TN-2-81/15
PUB DATE 30 Sep 81
CONTRACT 400-80-0108
NOTE 48p.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Beginning Reading; Elementary Education; *Language Processing; *Listening Comprehension; *Oral Language; Psycholinguistics; *Reading Comprehension; Reading Instruction; Sentence Structure; Speech Communication; Structural Analysis (Linguistics); *Suprasegmentals

ABSTRACT

Noting that the lack of prosodic information in printed text may be a source of difficulty for children who are learning to read, this paper explores the features of language underlying the acoustic and perceptual segmentation of sentences into meaningful units. Using evidence from studies in speech production and perception, the paper addresses two issues: (1) What principles of sentence organization should guide the physical segmentation of text as an aid for poor readers? and (2) What intrasentence units, if made explicit, would best facilitate children's induction of effective reading comprehension strategies? Throughout this inquiry, two types of sentence organization are discussed--syntactic organization, and information processing organization. The first part of the paper examines the nature of speech units by looking at sentence perception, and describes some assumptions of comprehension processes that serve as a conceptual framework for later discussion of several psycholinguistic approaches to the study of immediate processing units in speech perception. The second part considers studies of speech production and describes the prosodic features of spoken sentences and their complex relationship to syntactic organization and information "packaging." (HOD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED208388

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

SOUTHWEST REGIONAL LABORATORY
TECHNICAL NOTE

DATE: September 30, 1981

NO: 2-81/15

SENTENCE PERCEPTION IN LISTENING AND READING

David P. Snow and James H. Coots

ABSTRACT

This paper examines the role of suprasegmental phonology in speech processing and its implications for reading instruction. Evidence for the hypothesis that prosodic features cue the boundaries of perceptually functional units (e.g., phrases) in spoken sentences, thus assisting the listener in the immediate segmentation of verbal information, is reviewed in detail. Findings from both comprehension and production studies suggest that the perceptual organization of sentences is guided by syntactic structure as well as by information-processing requirements related to meaning and memory capacity. It is argued that prosodic cues are not well represented in written text, a fact which may underlie the difficulty that many children experience when learning how to comprehend what they read. The use of phrasally segmented text accompanied by adult modeling of prosodic rendering is suggested as a means for facilitating the child's induction of organization strategies during reading.

This report was prepared under Contract No. 400-80-0108 with the National Institute of Education, Department of Education. Its contents do not necessarily reflect the views of the National Institute of Education or of any other agency of the United States Government.

This document is intended for internal staff distribution and use. Permission to reprint or quote from this working document, wholly or in part, should be obtained from SWRL, 4665 Lampson Avenue, Los Alamitos, California 90720

CS806343

SENTENCE PERCEPTION IN LISTENING AND READING

David P. Snow and James H. Coots

Academic instruction is guided by a task-analysis which builds on the abilities students already possess and focuses on the skills they need to learn. This sequence of instruction, which proceeds from acquired skills to new skills, is as important in learning to read as it is in more substantive areas of the curriculum. Children typically begin reading instruction with very advanced oral language skills. Even very young readers who experience difficulty with reading comprehension usually are able to demonstrate proficiency in oral language comprehension (Guthrie & Tyler, 1976; Oaken, Wiener, & Cromer, 1971). Since most school children have already acquired a rich variety of oral language comprehension strategies prior to formal reading instruction, it is usually not necessary to teach them language comprehension skills per se, but rather to focus on differences between comprehension processes in listening and those required for reading. Thus, reading instruction is guided by an implicit analysis of the differences between oral language and reading.

The most salient difference between auditory and graphic modalities is, of course, the different media for the representation of words. Reading instruction in the early grades therefore emphasizes decoding skills, which are required for translating graphic segmental symbols into units of speech. This instruction consists of teaching print-sound correspondences paralleling the sound-meaning correspondences that children already know.

Although differences in segmental and lexical representation are important in the distinction between speech and writing, there are many

other differences as well, any one of which may represent an obstacle for children in their efforts to bridge the gap between oral language competence and reading comprehension. Among these additional differences between speech and writing are a) availability of contextual information, b) opportunity for interaction between speaker and listener, and c) degree of spatial and temporal commonality with the speaker (Rubin, 1978). Perhaps the most pervasive difference between modalities is that speech contains prosodic features (stress, intonation, and timing) which are not systematically represented in written material. Although punctuation, such as the use of commas, captures some prosodic features, there is little systematic correspondence between sound and symbol in suprasegmental phonology. The lack of a one-to-one correspondence between punctuation symbols and their acoustic correlates is perhaps a more serious problem for the beginning reader than the lack of one-to-one print-sound correspondences on the level of segmental phonology.

This suprasegmental difference between speech and writing is important because prosodic features give the listener cues about the meaning and structure of utterances. Prosody does this in at least three ways. First, stress contours signal differences between given and new information, and thus contribute to cohesion. Second, prosody segments utterances in conceptually unified groups of words which can be processed efficiently. And finally, stress, pitch, and timing provide cues to the hierarchical parsing structure of sentences, which is a more refined organization of prose than that involved in conceptual segmentation. These functions of prosody all assist the listener in the task of imposing organization on sentence strings.

The lack of prosodic information in printed text may be a source of difficulty for children who are learning to read. In fact, poor readers typically fail to organize text into meaningful units (Golinkoff, 1975-1976). This difficulty with text organization implies that while some children (good readers) are able to cope with the absence of prosodic information in text, others are unable to organize text without specific training.

Poor readers might benefit from cues in the text which assist them in text organization. One type of cue which is the object of current research efforts at SWRL consists of segmenting sentences into meaningful phrases and clauses--that is, into groups of words that "go together"--as described by Coots & Snow (1981). This representation of text captures some of the natural marking of sentences that is given auditorily through intonation, timing, and stress.

This paper explores the features of language underlying the acoustic and perceptual segmentation of sentences into meaningful units. Using evidence from studies in speech production and perception, we address such questions as 1) What principles of sentence organization should guide the physical segmentation of text as an instructional aid for poor readers? 2) What intrasentence units, if made explicit, would best facilitate children's induction of effective reading comprehension strategies?

Throughout this inquiry, two types of sentence organization are discussed: syntactic organization and information-processing organization. These two dimensions can serve as labels for alternative hypotheses concerning the ideal unit for phrasally segmented text. The syntactic segmentation of sentences assigns form-class labels to words, and

hierarchical levels of constituent boundaries to groups of words. The analysis leads to a parsing of each sentence; that is, to an elaborate description of surface syntactic relations analogous to a labelled bracketing (e.g., Chomsky & Halle, 1968). If phrasally segmented text is constructed so as to function as a cue for sentence parsing, the segmentation would reflect major syntactic units such as noun phrase and verb phrase.

The other type of sentence analysis groups words into blocks of information as a first step in the global organization of sentence strings. This type of analysis also takes constituent structure into account, but it is additionally sensitive to information-processing characteristics of sentence perception, such as length and meaning of units. Segmentation based on information characteristics would result in units reflecting constraints on length, meaningfulness, and intention, in addition to syntactic structure.

Experimental and observational evidence suggests that speech is perceptually organized in groups of words defined by both linguistic and information-processing parameters. Since these units seem to play a major role in global sentence-recognition processes, they suggest a basis for text segmentation that might be most helpful for young readers grappling with the problem of imposing meaning on complex reading material.

Part I of this paper examines the nature of speech units by looking at sentence perception. The first section of Part I describes some assumptions about comprehension processes that serve as a conceptual framework for later discussion of several psycholinguistic approaches to the study of immediate processing units in speech perception. Part II

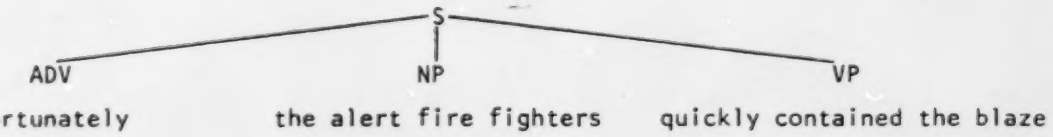
considers studies of speech production. It describes the prosodic features of spoken sentences and their complex relationship to syntactic organization and information "packaging."

PART I

Constituents in Immediate Speech Processing

Sentences differ from unordered lists of words or digits by virtue of having an internal organization. This organization is reflected in the constituent structure of sentences, which formally defines the grammatical relations existing between individual words, phrases, and larger units. In this paper, the primary concern is with the major immediate constituents of sentences and their role in sentence perception. These include subject noun phrases (NP), verb phrases (VP), and sentence adverbs (ADV).[1] These are major syntactic units because they are not subordinate to any other category symbols except for the sentence symbol (S) itself.

An example of the relationship among major constituents is shown below. Each major constituent can be further analyzed into more detailed



[1] Sentence adverbs modify the sentence as a whole and often occur in a sentence-initial position. They include a large set of expressions like naturally, in all probability, evidently, as well as sentence-initial clauses like having lost the game, by going to Cincinnati, etc. (Jackendoff, 1972). Other types of adverbs describe events temporally or spatially (e.g., at the park, during the summer). Like sentence adverbs, these are sometimes treated as constituents of the sentence rather than of the verb phrase (e.g., Bach, 1974). In this paper, these phrases as well as all subordinate clauses are described as "higher-level" adverbs; that is, they are subordinate only to the S node.

constituents. For example, the verb phrase contains a verb (contained), an object noun phrase (the blaze), and a manner adverb (quickly). This type of surface structure analysis can be continued until each word is assigned a form class which specifies its grammatical relation to other words within the same immediate constituent. The complete analysis is represented by a labelled bracketing, as shown below.[2]

[[Fortunately] ADV [[the] D [alert] A [[fire] N [fighters] N] N] NP
[[quickly] ADV [contained] V [[the] D [blaze] N] NP] VP] S

The importance of this analysis is that it shows the grouping of words in terms of grammatical relations, such as the relation between subject and predicate (NP and VP within S) or main verb and object (V and NP within VP). Since these syntactic relations are closely tied to the semantic content of the sentence, it is assumed that sentence comprehension involves a stage of surface syntactic analysis whose outcome is analogous to the labelled bracketing shown above. This stage of sentence perception is referred to as the parsing of a sentence string.

Although the psychological processes and strategies underlying the parsing of sentences are not well understood, it is generally agreed that the analysis of sentences is not carried out all at once, especially if sentences are long and complex. The study of comprehension processes suggests that listeners analyze sentences in chunks of information, that is, in groups of words, and that the boundaries of these chunks coincide with major phrase and clause boundaries.

---

[2]Symbol Key: A = adjective, ADV = adverb, D = determiner, N = noun, NP = noun phrase, V = verb, VP = verb phrase, S = sentence.

Evidence supporting this view of sentence processing comes from studies in immediate memory for sentences (Jarvella, 1971) and from studies of segmented text, such as those carried out by Anglin and Miller (1968) and Graf and Torrey (1966). Graf and Torrey, for example, presented sentences in whole phrases and clauses, the boundaries of which are shown by slash marks (/) in the following example.

(1) During World War II / even fantastic schemes / received consideration / if they gave promise / of shortening the conflict.

When subjects read sentences presented in phrasal and clausal units, as above, comprehension and recall were better than when they read sentences in other types of arbitrary groupings, such as in (2).

(2) During World War / II even fantastic / schemes received / consideration if they gave / promise of shortening the / conflict.

This finding cannot be explained by appeal to the well-known psychological strategy of chunking information in order to facilitate memory retention. If this were true, presentations (1) and (2) ought to work equally well, since both presentations segment the to-be-remembered sentence into the same number of chunks. The most plausible explanation for the facilitative effect of segmentation by major grammatical units is that this format "packages" information in a way that makes it easier to process; that is, it segments sentences into natural perceptual units.

The relation between organization and perception is exemplified in studies of the amount of verbal material adults can perceive and remember from brief visual exposures (up to 250 msec.) which correspond to a single

eye fixation, or one brief glance. Commenting on these studies, Smith (1978) concluded that the amount of material that can be processed in a single fixation depends on its organization. If the stimulus string is an unordered list of letters, subjects can only perceive about four or five letters at a time. If letters are organized into random strings of words, perceptual capacity increases to about two or three words (up to 12 letters). When words are further organized into short, meaningful clauses, subjects can perceive in a single moment about five words (or 25 letters). These results are illustrated below in Figure 1.

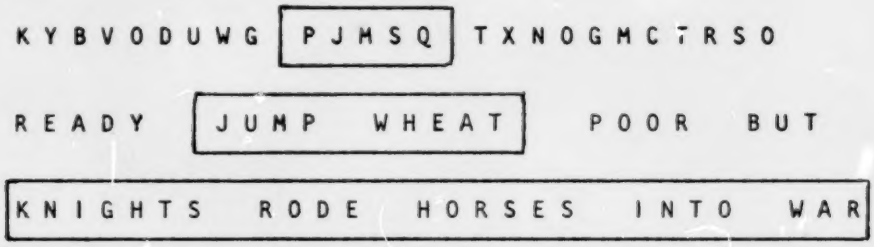


Figure 1. Information perceived in a single glance, as a function of the stimulus organization (from Smith, 1978).

The amount of information perceivable in a single exposure probably approximates the capacity of short-term storage without rehearsal. The findings cited by Smith (1978) have two important implications. First, they show that immediate processing capacities for verbal material vary according to the syntactic-semantic organization of the material. Second, they suggest the importance of phrasal length in visual sentence perception: units of perception seem to be in the neighborhood of five words in length.

Studies using phrasal text (e.g., Graf & Torrey, 1966) as well as the above findings cited by Smith (1978) show that phrasal organization facilitates the perception and recall of verbal material. One explanation

for this facilitation is that phrasal encoding helps the reader to more rapidly arrive at a correct sentence parsing. In this case, we would expect phrasal units to be syntactic in nature, corresponding to major constituents such as NP, VP, ADV. However, this is not always true. For example, the phrase if they gave promise in (1) is not a major constituent, and Knights rode horses into war (Figure 1) is an entire sentence. Another explanation, which is explored below, makes appeal to other aspects of sentence processing such as meaning and memory.

A Processing Model

Using an information-processing model of speech perception (e.g., Jarvella, 1971; Kintsch & Vipond, 1979), we assume that sentences are initially segmented in working memory in groups of lexical items forming clauses or phrases. This surface structure string is interpreted through a series of phonological, lexical, surface syntactic and deep syntactic analyses in short-term storage (STS), eventually resulting in an abstract semantic representation. The semantic (propositional) representation is transferred to long-term storage for more durable retention, while the verbatim lexical representation rapidly fades from STS.

The ideal unit of surface analysis is a major phrase or clause because these units satisfy two requirements of a short-term store: 1) they correspond to a meaningful grouping capable of semantic representation, and 2) they are small enough not to overwhelm the limited capacity of STS. In sum, grouping words into constituent units is important in comprehension because it facilitates the transfer of information from surface to semantic representations via the limited capacity of a short term memory processor (Fodor, Bever, & Garrett, 1974). Following Fodor

et al., we refer to this stage of analysis as the immediate perceptual segmentation of sentences.

One objective of this paper is to explore the characteristics of the immediate units of speech processing. What syntactic and semantic features characterize groups of words that "go together"? The processing model sketched above suggests that groups of words that go together are linked by semantic relations, i.e., they express "ideas." The most basic ideas, such as occur in almost every sentence, specify a) the roles of participants or objects in relation to a single action or state (e.g., the man read the newspaper, the child slept, the actor felt nervous), or b) the relation existing between an entity and an attribute (the dancer was graceful) or an orientation such as location (toys were on the floor). All of the above ideas are expressed syntactically as simple sentences. The semantic relations between words are expressed by grammatical relations between constituents such as subject, verb, object, or a simple predicate complement.

Semantic relations between lexical items can also be expressed by embedding phrases or clauses within a sentence constituent such as a subject noun phrase (e.g., the man reading the paper, the sleeping child, the toys on the floor). Although these complex phrases are not grammatically independent units, they express ideas in the same way that complete simple sentences do. However, the grammatical relations are different: In complex nominal phrases, the underlying content is expressed by the grammatical relation between noun and adjunct and/or between modifier and noun.

The fact that constituents may be grammatically complex permits complex sentences to express several interconnected ideas at once. An important implication for sentence processing is that syntactic units such as noun phrases and adverbs may function conceptually like independent simple sentences, depending on the complexity of content expressed. The notion that listeners are sensitive to the ideas, or information structure, of sentences is consistent with the findings of several psycholinguistic studies of the perceptual segmentation of speech. Two examples of these studies are reviewed next. They illustrate a major theme of this paper: that units of language are not determined by strict parsing boundaries (e.g., the syntactic border between subject and predicate) but by the boundaries of groups of lexical items having immediate syntactic-semantic relations to one another.

The Psychological Reality of Constituents

The conceptual unity of phrases as a function of constituent relationships has been shown experimentally by probe-latency studies. One example is the "two-word" probe technique, which is a recognition paradigm. Subjects are given a two-word sequence subsequent to exposure to a stimulus sentence. The task is to judge whether or not the two-word sequence occurred in the previously heard sentence. If the test sequence occurred within a constituent, it should be recognized more quickly than if the probe crossed a constituent boundary between different conceptual sequences of words. Thus, latency of response gives some indication of subjects' use of grouping strategies in sentence comprehension.

An interesting study using the two-word probe technique was reported by Stewart and Gough (1967, cited by Fodor et al., 1974). This study,

investigating the effects of surface constituent structure on sentence processing, used sentences such as (3) and (4) which controlled for serial position and lexical effects.

(3) [The presidents of large corporations] NP [pay millions of dollars in taxes each year] VP

(4) [When profits are large] ADV [corporations] NP [pay millions of dollars in taxes each year] VP

A two-word test probe such as large corporations is contained within a major unit in sentence (3) (it is part of the noun phrase), but it crosses a major boundary in sentence (4). If major constituents are units in the psychological processing of sentences, the adjacency of the test words should be more difficult to recognize in (4) than (3). The latency results confirmed this prediction.

Other probe pairs give even more insight into the segmentation of these sentences. Stewart and Gough (1967) also used probes such as corporations pay. In this case, the word pair crosses the same major boundary in each sentence, the NP-VP boundary. Thus, the surface-structure constituency is the same. However, response latencies in this case were longer for (3) than for (4). As suggested by Fodor et al., this asymmetry in response latency reflects the difference in grammatical relations between corporations and pay. In (3), the noun corporations is grammatically related to the subject noun head presidents. It is therefore grouped with the entire subject noun phrase. In sentence (4), corporations also borders on the NP-VP boundary, but it is perceptually grouped with the verb because of the subject-verb relation between them, which does not exist in (3). In sum, people seem to group words in psychological units that maintain immediate grammatical relationships.

The same finding has also been reported in studies using the well-known "click paradigm." In this technique, subjects are asked to locate the position of a short burst of noise (click) within a stimulus string such as a sentence or word list. The perceived locus of an interrupting noise is often different from its objective location--a phenomenon called subjective displacement. Ladefoged and Broadbent (1960) found that errors in subjects' perception of click location were larger and more frequent when clicks were superimposed on sentences than on nonmeaningful strings of words. A number of subsequent studies showed that click displacement for sentential material was associated with the major constituent boundaries in sentences, especially finite clausal boundaries. That is, subjects tended to subjectively displace auditory interruptions toward the boundaries of major constituent units in sentences (Fodor & Bever, 1965).

Click displacement is usually explained in terms of the concept of "perceptual closure"--the tendency to resist interruption of the integrity of perceptual units. The results bear on sentence processing by showing that sentences are perceptually organized in constituent units.

In one study showing the click displacement effect (Garrett, Bever, & Fodor, 1966), the possible effects of prosody were controlled by creating pairs of structurally different sentences from spliced tape segments. This procedure ensured that the critical constituent boundaries were acoustically identical. An example of the materials is shown below, with arrows indicating the bounardies which "attracted" click displacements and a dot (•) showing one of the objective click locations.

(5) [The retiring chairman whose methods still greatly influence the company] NP [was given an award] VP
↑

(6) [As a direct result of their new invention's influence] ADV [the company] NP [was given an award] VP
↑ ↑

Subjects tended to subjectively displace the locus of click interruptions to major constituent boundaries in each sentence. In (5) this boundary coincided with the NP-VP (subject-predicate) break. However, in (6) clicks were attracted to the adverbial constituent boundary. This means that the subject-predicate break was sometimes effective in attracting clicks (sentence 5) and sometimes not[3] (sentence 6). As Fodor, Bever, and Garrett (1974) concluded, "If the primary perceptual segmentation of a sentence depends solely on its surface tree, it is hard to see why there should have been this difference" (p. 337).

These results parallel the findings of probe-latency studies as exemplified in sentences (3) and (4). Both experimental paradigms suggest that listeners group words into perceptual units on the basis of immediate syntactic-semantic relations between words. Thus, for example, the noun phrase the company is perceptually grouped with the verb phrase in (6) but not in (5), because these constituents have an immediate relation in the former case but not in the latter.

---

[3]Fodor et al. note also that the displacement effect is more consistent with sentences like (5) than in (6). They interpret this finding as support for the hypothesis that the primary units of speech processing correspond to the surface structure reflexes of deep sentoids (underlying sentences in linguistic theory). In this paper, we have argued that perceptual units correspond to groups of words expressing coherent semantic and syntactic relationships (which is consistent with the deep sentoid theory), but we have not insisted that such units correspond to underlying sentences in linguistic theory. For an interesting discussion of this issue, see Carroll and Tanenhaus (1978).

Summary of Part I

The constituent structure of sentences is important in comprehension because it specifies the syntactic relationships which are necessary for correct sentence interpretation. Parsing is a stage of sentence processing in which all grammatical relations are tacitly identified. A distinction is made between parsing and an initial processing step which segments sentence strings into larger conceptual units or "packages." This segmentation process is accounted for in an information-processing model of comprehension which posits at least two different memory representations for sentences, verbatim and semantic, as well as distinct memory stores. The psychological segmentation of sentences facilitates the transfer of information from a verbatim surface string to an abstract semantic representation via the limited capacities of short-term storage. Efficient processing units are short (about five words in length), meaningful phrases or clauses. "Meaningfulness" implies that the lexical constituents of phrases are grouped together by immediate grammatical relations. Phrases in speech perception may therefore cross major syntactic boundaries if the words on either side of the boundary are linked by a grammatical (hence, semantic) relation.

Part II continues our inquiry by exploring the prosodic systems of speech and their relation to the organization of sentences. Among other functions, prosody seems to be associated with the segmentation of speech into conceptual units. This function of suprasegmental phonology is of particular interest in the context of reading comprehension instruction. In Part I, it was suggested that the units of speech processing express ideas--"packages" of the content of sentences. To the extent that prosodic

features mark the boundaries of these units, prosody can be viewed as a dimension of linguistic form that contributes to the expression of meaning. However, this sound-meaning correspondence is lost in the translation from spoken to written language.

These observations imply a serious problem for the young reader: Speech prosody depends on meaning, but neither prosody nor meaning is directly represented in writing. Not unexpectedly, poor readers experience difficulty with both the sound and the meaning of written text. On the level of comprehension, they fail to group words into meaningful units and therefore respond poorly to questions requiring a deep semantic processing of the text. In oral reading, poor readers tend to read with little variation in stress, pitch and timing; that is, they read with little "expression." Although many factors may interfere with children's reading comprehension skills (e.g., deficits in vocabulary or decoding skills), it is reasonable to assume that one adverse factor is the loss of prosodic information in written text. In order to plan instruction for helping children to organize text without the benefit of prosodic cues, we need to know more about the function of organizational signals that are present in speech, but missing in writing. What does it mean to read with "expression"? And how is this related to comprehension? With these questions in mind, we turn next to a consideration of speech prosody.

PART II

Prosodic Features

Careful study of the sound and cadence of speech reveals that oral communication is rich in features analogous to those found in music:

variations in loudness, pitch, and tempo. The interplay of these features gives a sensation of pattern and rhythm. This suprasegmental accompaniment to speech serves at least three functions: a) it indicates the speaker's attitude and affect; b) it focuses attention on particular aspects of the message, such as distinguishing between new and given information; and c) it marks the structural boundaries of sentence and intrasentence units. The last function is of special interest in this paper since it is directly related to the basic processes of comprehension discussed in Part I.

Prosody refers to acoustic patterns in speech produced by variations in fundamental frequency, amplitude, and timing which do not contribute directly to the identification of phonemes (Crystal, 1969; Larkey, 1979). These features are often called "suprasegmentals" because of their independence from the identifying parameters of phonetic segments. In fact, suprasegmental features often apply to units larger than segmental phonemes; that is, they may apply to syllables, words, and higher-order units of structure.

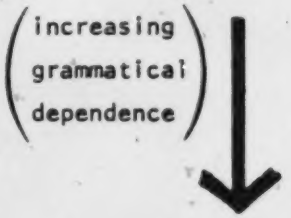
The acoustic parameters of fundamental frequency, amplitude, and timing are psychologically perceived as pitch, loudness, and duration, respectively. Variations of these features in spoken sentences lead to perceptions of intonation, stress, and rhythm. Intonation refers to the pattern of pitch alterations in sentences. Stress is usually regarded as the relative perceived loudness of syllables. And the different patterns of stress and timing lead to perceptions of rhythmicality. The main perceptual parameters of prosody, then, are pitch (or intonation), stress (or loudness), and duration (or timing or rhythm).

The best suprasegmental sources of information regarding syntactic boundaries are the temporal structure of spoken sentences and their patterns of fundamental frequency (Larkey, 1979). Both of these acoustic parameters are closely related to the complex feature of stress, which is discussed in more detail following consideration of the temporal features of speech.

Temporal Structure

Many studies have used subjects' judgments of pauses in speech as the basis for subjective segmentation of sentences into meaningful groups of words. Research of this type implicates the relationship between temporal phenomena of speech and major structural boundaries as important in speech comprehension. The typical finding has been that syntactic boundaries between sentences, clauses, and phrases are marked by a lengthening of pauses, terminal segments, or both.

Goldman-Eisler (1972) has presented evidence indicating that the frequency and duration of pauses in speech is related to the grammatical characteristics of the boundaries at which pauses occur. Studying the pausing of speakers during reading and spontaneous speech, she attempted to correlate pausing characteristics with the following types of structure:



sentences
coordinate sentences
subordinate clauses
relative clauses
words within clauses

Goldman-Eisler described these structures in terms of a gradient of grammatical dependence, as in the order listed above. Sentences are the most independent units, while individual words within clauses have the greatest degree of dependence on adjacent units (words). Using a fairly long pause duration as the criterion for nonfluency or hesitations (pause = 500 msec), Goldman-Eisler found that the mean frequencies of hesitations at the specified boundaries reflected the degree of grammatical dependence between bounded units. That is, hesitations of 500 msec or greater were most frequent between sentences, less frequent between clauses, and were very infrequent between words within clauses. This finding supports the notion that speech is organized into conceptually coherent units corresponding to sentences (and to a lesser extent, clauses), and that the semantic integrity of speech units is reflected in the degree of temporal integration across them.

In these and other performance studies of pausing, considerable variation in pause duration has been found at all structural boundaries except for sentence boundaries, where pauses in reading tend to cluster about a mean length of 1.00 to 1.25 seconds. This variation suggests that pauses are at best an optional marker of structure for intrasentence units. Perhaps the best way to describe the function of pauses is to refer to their role in perception rather than production: When speakers do pause, the increasing duration of the pause increases the likelihood that listeners will perceive a major syntactic boundary (Larkey, 1979) or complete a perceptual unit in speech processing (Bever, Lackner, and Kirk, 1969).

One of the difficulties in analyzing the performance structure of pauses is that many hesitations in speech reflect planning time. As such, they reveal more about strategies for sentence production than for comprehension. In reviewing the literature on pausing phenomena, Clark and Clark (1977) concluded that there are three major places where speakers tend to pause.

a) Grammatical junctures, especially sentences. These are the longest and most frequent pauses. They are typically silent pauses. Sentence boundaries appear to be the most consistent locus of structurally motivated pausing.

b) Other constituent boundaries, including phrasal boundaries. Usually filled pauses, these seem to represent planning hesitations and suggest a model of speech production which proceeds constituent by constituent.

c) Before the first content word within constituents (Boomer, 1965). This is a hesitation pause indicating decisions about planning and coding the utterance. Such pauses are either silent or they involve repetitions which retrace the utterance back to the beginning of the constituent, e.g., "the/the dirty cups." This type of pause does not mark grammatical structure but reflects the speaker's planning of sentence execution in constituent units. The structure of repetitions and corrections also shows that speakers attempt to deliver the constituent as a single phonological and perhaps conceptual unit.

In sum, pausing is unreliable as a structural cue because hesitations are optional and they may occur at locations other than major boundaries.

Another dimension of speech timing is segmental lengthening, a measure that reflects the relative duration of phonetic segments. The last stressed syllable of phrases and higher-order units is slightly prolonged, resulting in a segmental duration for that syllable that is longer than its characteristic timing. Acoustic studies of speech indicate that constituent-final lengthening occurs more frequently and consistently than pauses. It is therefore a prosodic feature of major interest in this discussion, which seeks to specify the reliable structural cues which are given by suprasegmental phonology.

Since most of the variance in vowel length is accounted for by the intrinsic durations of differing vowel segments (Klatt, 1975), most studies of the relationship between segmental lengthening and structure have used sentences that are structurally ambiguous; that is, different syntactic bracketings can be assigned to the same string of segments. Using materials of this kind, Cooper, Paccia, and Lapointe (1978) found evidence in support of the claim that the duration of segmental lengthening is proportional to the height of constituents in a tree-structure (hierarchical) representation. Thus, phrases coterminous with major branches high in the structure are marked by greater final-segment lengthening than those that are embedded at lower levels in the hierarchical arrangement.

Cooper et al. (1978) asked adult subjects to read sentences such as (7), (8), and (9) below. The structurally ambiguous constituent in each case is underlined.

(7) My Uncle Abraham presented his talk naturally.

(8) Here is the famous Duke James.

(9) Pam asked the cop who Jack confronted.

The structural ambiguity of these sentences arises from the fact that the underlined constituents can be parsed at differing heights in the tree structure. The adverb naturally in (7) can be either a sentence adverb (dominated by the S node) or a manner adverb (dominated by the VP node). This can also be expressed notationally by alternative labeled bracketings, as shown below.

(10) [My Uncle Abraham]NP [presented [his talk]NP]VP [naturally]ADV

(11) [My Uncle Abraham]NP [presented [his talk]NP [naturally]ADV]VP

In (10), the noun talk completes both a lower-level noun phrase and a higher-order verb phrase. In (11), the same noun completes only a noun phrase. The alternative labeled bracketings provide a notation for describing the concept of immediate grammatical relations as discussed in Part I. Significant surface structure breaks such as in (10) are characterized by the absence of a grammatical relation between constituents, and such boundaries tend to be notationally signaled by the presence of multiple closing parentheses in a labeled bracketing.

The Cooper et al. (1978) subjects read sentences such as those in (7) through (9) after being given cues as to how to interpret the ambiguous strings. For example, in (8), one reading would be cued by "James meets the Duke." The relative length of the key phrase-final syllable (talk, Duke, cop) was consistently longer in the reading in which the key word completed a higher-level constituent. For instance, the vowel segment in

talk was 15% longer in the context of reading (10) than in the different structural context of (11).

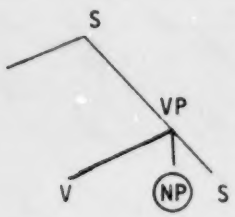
These results suggest that speech coding is sensitive to the "height" of phrases in the surface structure. Cooper et al. (1978) concluded that phrase-final segments are differentially lengthened in proportion to their hierarchical level in the surface structure. This is illustrated with the following two sketches of alternative constituent structure.

The phrase-final segment of the circled NP will be marked by greater duration in (a) than in (b), because this constituent completes a phrasal unit higher in the surface structure in (a) than in (b). The proportional lengthening could be accounted for by a psychological model that cyclically assigns prosodic markers to phrases at each level of structure, as discussed later in this paper in remarks on stress assignment. If all phrasal boundaries are marked by lengthening, the phrase-final segment of the circled NP above will be lengthened more in (a) than in (b) because it receives the accumulation of durational values for both NP and VP.

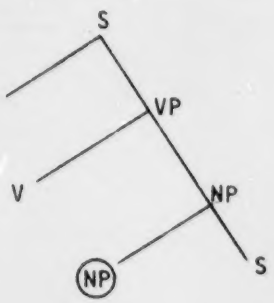
This "simultaneous phrase" principle, however, cannot account for the durational values which differentiated the two readings of sentence (9). In one interpretation of sentence (9), the ambiguous embedded

sentence who Jack confronted functions as an indirect question, which is among the class of predicate complement constructions. In the other reading, the embedding functions as a relative clause. These two representations are sketched below.

(a) predicate complement



(b) relative clause



In the Cooper et al. (1978) study, there was a more pronounced temporal marking of the circled NP constituent in (a) than in (b). The prosodic difference is attributable to the different height of the key constituent in the tree structure, but not to an effect based on multiple phrase boundaries. Thus, constituent-final lengthening seems to be sensitive mainly to the height of constituents in a hierarchical representation, with "height" determined by the number of nodes which dominate a given constituent.

Cooper et al. (1978) also measured pause durations between the key constituents of structurally ambiguous sentences. The data showed that there were longer pauses following the phrase-final key words when the subsequent constituent was attached to a higher sentence node. Thus, pauses showed an increase in duration at the same places where relatively greater segmental lengthening was also observed. Although the mean pause

differences tended to be of greater magnitude than the differentials in segment length, the pause data showed much more intersubject variation as reflected in larger standard deviations. This kind of variation suggests that pausing is an optional and inconsistent prosodic marker of structural boundaries within sentences.

Given the variability that characterizes pausing, the most reliable temporal measure of speech coding seems to be the timing of interstress intervals (ISI's), which sums the variable effects of both segmental lengthening and pauses. Interstress intervals reflect the time duration between the onsets of stressed syllables. Interval lengths vary according to two major factors: a) the number of intervening unstressed syllables, and b) syntactic boundaries. In spite of the variability introduced by these factors, the time interval between stressed vowels tends to cluster around a mean value of .4 to .5 seconds (Lea, 1975). This tendency toward isochronous ISI's supports the claim that English is a stress-timed language, and suggests that prominent variations or breaks in this regular temporal pattern are likely to function as salient cues to the structural pattern of sentences. Lea found, for example, that ISI's crossing clausal boundaries were about twice as long as the mean value, and about three times as long when they cross sentence boundaries. As described by Larkey, "It was as if the speaker waited an extra beat at clause boundaries and an extra two beats at sentence boundaries" (p. 33).

Lea (1975) also observed ISI variations of a lesser magnitude across constituents within clauses. In general, Lea's findings support the conclusions of Cooper et al. (1978) that temporal effects are proportional

to the hierarchical position of constituents in the surface syntactic analysis.

An interesting aspect of Lea's work on speech timing is that ISI's reflect time durations between stress beats. The consistent data obtained by this measure suggest a connection between temporal effects and stress placement in regard to the prosodic marking of sentence structure. This connection is further explored when the topic of stress placement is discussed, below.

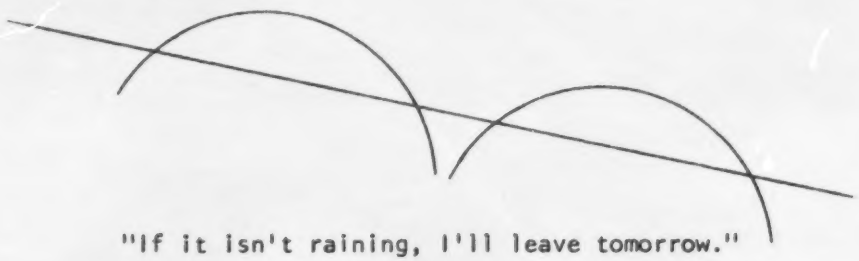
Fundamental Frequency

A spoken sentence is accompanied by constant changes in the fundamental frequency (F_o) of the speaker's voice. The overall pattern of these pitch variations describes an intonation contour or characteristic "melody" of sentences. A number of linguists have regarded intonation contour as a prominent prosodic marker of the natural units of language. Groups of words spoken under a single intonation contour present information in unified packages of content variously called information blocks (Grimes, 1975), information units (Halliday, 1967), or tone units (Crystal, 1969). F_o contours thus represent an important cue for the initial perceptual segmentation of sentences. The present section explores the structure of intonation and seeks to specify some of the syntactic and semantic features of intonation boundaries.

The global intonation contour of sentences normally describes a rising and falling frequency curve.[4] F_o rises sharply at the first

[4] A major exception to this pattern is the contour of yes/no questions.

syllable of an utterance and declines gradually throughout the sentence until the final stressed syllable, where a more rapid fall in pitch occurs. In complex sentences, this archetype F_0 contour (Lieberman, 1967) is typically composed of smaller local intonation curves corresponding to major syntactic boundaries such as clauses. The sketch in Figure 2 shows an idealized example of local contours which envelope each clause of a single sentence. The archetype contour is an idealized, abstract entity corresponding to the curvilinear mean of the changing frequency signal. The sentence as a whole also has a changing mean frequency, which could be represented by a global archetype contour spanning across the local contours, but which is shown more simply as a straight line in Figure 2.



"If it isn't raining, I'll leave tomorrow."

Figure 2. Idealized sketch of local F_0 contours.

Local contours within sentences are characterized by a fall in pitch at major constituent boundaries and a subsequent rise in pitch on the first stressed syllable following the boundary. Studies of the grammatical structure underlying intonation suggest that fall-rise patterns in F_0 are fairly reliable prosodic markers of the boundaries of main clauses and even certain phrasal units. For example, Cooper and Sorensen (1977) compared fall-rise patterns at the clause boundary in conjoined and embedded structures. The structure of these two sentence types is sketched below, showing clause conjunction in (12a) and embedding in (12b):

(12) (a) [main clause]S [main clause]S

(b) [main clause [embedded clause]S]S

Examples of the sentences used to represent these structures are shown in (13) and (14). Cooper and Sorensen studied the magnitude of fall-rise changes in F_0 at the major syntactic boundaries, which are marked with a slash (/) in the examples.

(13) (a) Anthony was surprised / and Raymond became upset.

(b) Anthony was surprised / Andréa became upset.

(14) (a) Marie was listening to the song / and Raymond was playing.

(b) Marie was listening to the song / Andréa was playing.

The materials controlled for phonetic environment by using pairs of sentences nearly matched in terms of segmental phonemes but differing in structure. F_0 measurements focused on critical regions before and after the major boundaries, for example, the syllable /praiz/ in surprised and the sequence /dre/ in and Raymond. All sentences showed an F_0 fall-rise contour across the syntactic boundary. This fall-rise pattern is characterized by an initial peak in F_0 at the onset of the pre-boundary stressed syllable. Pitch drops markedly through the course of this syllable's production, describing an F_0 valley at the end of the syllable. The post-boundary syllable shows a subsequent rise in pitch, resulting in another peak but one that is lower than the peak preceding the boundary.

Although Cooper and Sorensen (1977) found significant fall-rise contours across the boundaries preceding both conjoined and embedded clauses, the magnitudes of fall in F_0, and especially of the subsequent

rise, were greater in conjoined structures (13a and 14a) than in embedded ones (13b and 14b). Thus, the F_o valley was somewhat less prominent when one clause was embedded within another than when two clauses were conjoined.

Cooper and Sorensen (1977) also investigated intonation contours at the boundaries of some phrasal units that are pertinent to the questions of segmentation explored in this paper. Of particular interest are a) subject noun phrases containing postnominal modifiers, and b) preposed adverbial phrases. An example of each of these is given below, again showing the relevant boundary with a slash mark.

(15) The owner of the park / shows gold to the children.

(16) At Brockton's city park / Cher scolded the children.

Measurements of the intonation contour at these phrasal boundaries also showed significant F_o valleys. As in the previous examples, the two types of structure were differentiated by the magnitude of the post-boundary rise in pitch, which was much greater for preposed structures such as (16) than for phrases having a normal SVO word order such as (15). It is interesting to note in this regard that the boundaries showing the most prominent F_o valleys in speech (that is, 13a, 14a, and 16) are the ones that are often marked by commas in their graphic representation.

These findings indicate that both clauses and major phrasal units are typically marked by variations in F_o across the boundary just as they also tend to be accompanied by lengthened interstress intervals. In the next section, we explore the complex feature of stress and its important relation to both F_o contours and interstress intervals.

Stress

Although stress is traditionally regarded as the perceptual analogue of relative vocal intensity, there is considerable evidence to suggest that it is a more complex function of F_o, duration, and intensity, with F_o being the most important of the three (Larkey, 1979). Perhaps the best measure of stress is the integration of energy over a given segment, a measure that incorporates F_o, duration, and intensity. The acoustic correlates of stress perceptions are a) a rise and fall of F_o above and below the archetype intonation contour, b) segmental lengthening, and c) increased vocal intensity.

The major rule of stress placement in English is the Nuclear Stress Rule whereby primary stress is assigned to the right-most lexical constituent of a phrase,[5] e.g., brown^{2} eyes^{1}, John's^{2} uncle^{1}, catch^{2} the train^{1}, write^{2} a letter^{1}. These rules apply cyclically to sentences, with the scope of each cycle defined by successive levels of constituent structure (Chomsky & Halle, 1968). This is illustrated below with a hypothetical bracketing of constituents in which W, X, Y, Z represent lexical items. The stress assignments begin with the innermost brackets (Y Z), then proceed to the next higher level (X Y Z), each time assigning primary stress to the right-most constituent of the phrase and weakening other stress levels.

Stress Assignments	Domain of Stress Rule
(W (X (Y Z)))	
2 1	(Y Z)
2 3 1	(X Y Z)
2 3 4 1	(W X Y Z)

[5] 1 = primary stress. Other numbers indicate successively lower (weaker) levels of stress.

Primary stress in this phrase structure will fall on the final lexical constituent, Z. As the branching complexity of the structure increases, the difference in stress assignments between Z and neighboring constituents also grows larger, presumably increasing the likelihood that primary stress on Z will be perceptually salient.

Since the acoustic correlates of stress are a rise and fall of F_0 above and below the archetype contour line, an increase in vocal energy or amplitude, and segmental lengthening, the boundary between major phrases such as Z and neighboring constituents will tend to be marked by a combination of all the suprasegmental features which have been discussed in this paper. Thus, stress placement plays a central role in specifying the prosodic cues of constituent boundaries in sentences.

Discussion of Prosodic Features

This cursory review of prosodic phenomena has emphasized two major acoustic cues: lengthening of interstress intervals (timing) and local F_0 contours (intonation). Both timing and intonation have been related to stress assignment rules which place stress on the last syllable of higher-order constituents. Empirical studies show that contrasts in constituent structure between pairs of phonetically similar sentences are differentiated in vocal production by the relative magnitude of the interstress interval and F_0 valley across the major constituent boundary. The magnitude of these acoustic features is greatest at boundaries of higher-level constituents such as independent clauses or preposed sentence adverbial constructions. To a lesser extent, prosody also marks the boundaries between main and embedded clauses such as complement

structures. And finally, prosodic cues are weakest across highly dependent constituents low in the structure, e.g., verb-object. These findings support the claim that prosodic signals at intrasentence boundaries are proportional to the height of constituents in the hierarchical structure.

One implication of this finding is that prosody can differentiate between alternative interpretations of ambiguous sentences. For example, the sentence Pam asked the cop / who Jack confronted has two different interpretations depending on whether the embedding is parsed as a verb phrase constituent (17, below) or as a relative clause (18).

(17) She asked: "Who did Jack confront?" (higher-level parsing)

(18) Jack confronted the cop. (lower-level parsing)

Lengthening of the interstress interval across the structurally ambiguous boundary increases the likelihood that listeners will perceive a major grammatical break at that boundary. Thus, a slight pause after the cop would bias interpretation toward the reading in (17). Since phrasally segmented text serves a function analogous to salient suprasegmental signals, the physical segmentation of text at ambiguous boundaries would also tend to bias readers towards an interpretation which assigns constituents to the higher level of two alternative parsings. For example, consider the following segmentation of this sentence.

Pam asked the cop

who Jack confronted.

This might bias readers towards perceiving the embedded clause as a complement structure (17) rather than a relative clause (18). A similar effect can be observed in the following example.

Anthony was surprised

Andréa became upset.

This segmentation could potentially lead readers to parse the second clause as a coordinate structure (high in the structure) rather than a complement embedded at a somewhat lower level in the sentence structure.

Some examples similar to these are discussed by Crystal (1969) in remarks on the grammatical functions of stress and intonation. Among other examples, Crystal cites the following pairs of ambiguous sentences which are differentiated by prosodic features. Slash marks indicate boundaries provided by timing and intonation cues.

(19)	(a)	/ even if he told me / I'd go /	Subordinate + main clause
	(b)	/ even if he told me I'd go /	Subordinate + complement
(20)	(a)	/ my brother / who's abroad /	Non restrictive relative clause
	(b)	/ my brother who's abroad /	Restrictive relative clause

In these cases, segmentation of the ambiguous strings would tend to function in the same way that prosodic cues do; i.e., it would suggest an interpetation corresponding to the reading in sentence (a) of each pair rather than (b). This shows that phrasal segmentation must be used carefully in potentially ambiguous environments.

The limited examples and studies reviewed in this paper do not permit us to evaluate the claim that phrase-final lengthening (or any other signal) is proportional to the structural height of all boundaries. These studies do suggest, at least, that prosody marks the boundaries of major phrasal constituents, and that the saliency of prosodic cues varies in part as a function of the height of units in the surface structure.

However, we would like to review some arguments suggesting that the syntactic feature of structural height is not the major variable affecting prosody. As argued in Part I, which discussed the psychological reality of speech units, it appears that the important structural variable underlying prosodic phenomena is the notion of syntactic-semantic relations between words. Although the grammatical dependence between adjacent words across boundaries tends to vary inversely with the height of boundaries, structural height is not always a reliable index of these relationships. The following sentences, taken from Cooper and Sorensen (1977), illustrate the importance of grammatical relationships in the intonation structure of sentences. The sentences are parsed to delineate major constituents, in relation to the boundary (/) where a significant fall-rise in F_0 was observed.

(21) [The owner of the park]NP / [shows gold to the children]VP

(22) [At Brockton's city park]ADV / [Cher]NP [scolded the children]VP

The intonation boundaries observed in these sentences show the same type of segmentation that was suggested earlier by studies in sentence perception (cf. sentence pairs (3), (4) and (5), (6)). As in the discussion of those sentences, we focus here on the potential break between subject and predicate (NP-VP boundary) as a case study of variables affecting the segmentation of sentences.

In (21), a fall-rise in F_0 occurs at the subject-predicate boundary, between park and shows. In (22), however, there is no evidence to indicate that a similar intonation break occurs at the same constituent boundary between Cher and scolded. This asymmetry in intonation boundaries shows that the height of constituents in the tree structure is not the sole

basis for perceptually functional units, since the height of the subject-predicate break is the same in both sentences.

The distribution of intonation breaks is accounted for by considering the functional grammatical relationships among words. In (22), the words Cher, scold, and children are related grammatically (subject-verb-object) and semantically (actor-action-patient). They form a natural phrasal unit under a single intonation contour. In (21), however, two types of grammatical relationships are expressed by the constituents functioning as subject and predicate. One of these is a subject-verb-object-indirect object relationship between owner, show, gold, children. The other is a relationship between owner and park, which is expressed grammatically by embedding a postnominal adjunct within the subject NP. These two distinct relationships tend to be grouped under separate intonation contours, giving a subject-predicate break in (21) but not (22). As a result, the boundaries of functional units are typically signalled by a distinct shift in grammatical surface structure relationships. For example, note the lack of an immediate relation between park and shows (21) and between park and Cher (22).

In summary, the observational data suggest that prosodic cues mark major constituent boundaries which reflect groupings of words having grammatical relations such as modifier-noun, noun-adjunct, subject-verb-object. Recalling the processing model discussed in Part I, we can reasonably conclude that prosody assists listeners in the task of organizing words into perceptual units in memory, precisely because prosody redundantly signals a break or a significant change in the series of relat'onships between words.

The preceding sketch of prosodic phenomena is admittedly narrow in scope, and it fails to address many perplexing questions regarding the use of prosody in oral communication. One of the important residual problems is the question of the relationship among stress, timing, and intonation. In this paper, these acoustic parameters have been discussed as though they were all interdependent. In fact, we have speculated that various prosodic phenomena are interrelated as different acoustic derivatives of phrase-final stress-placement rules. However, a number of arguments suggest that this view is far too simplistic. For one, stress is sometimes quite independent of phrase structure. For example, stress may be used a) contrastively, b) to focus on new information, c) for emphasis, or d) to signal marked uses of pronouns, questions, and other structures.

Another observation more pertinent to this discussion is that intonation is more variable than timing features (Larkey, 1979), suggesting that intonation contours may be somewhat independent of phenomena such as constituent-final lengthening. In order to account for these observations, we speculate that timing may reflect the syntactic phrase structure of sentences, whereas intonation is used in a more variable manner to group words into perceptual units appropriate for short-term memory capacity. If intonation contours are viewed as being melodic patterns superimposed on the temporal structure of spoken sentences, then a significant fall or rise in F_0 at processing boundaries could serve to amplify the perceptual salience of simultaneously occurring features such as stress and rhythm. In this sense, different prosodic systems converge at local contour boundaries. Thus, the boundaries of speech units having

psychological significance are acoustically marked by several prosodic features occurring together. As an example, sentence (23) shows how clausal boundaries are accompanied by an interplay of a) stress (upper-case letters), b) segmental lengthening (diacritic bar), c) a possible slight pause (dash), and b) a fall-rise in F_0 (↓ ↑).

(23) If John should CĀLL ↓ -- ↑ please give him this MĒSSAGE ↓

The Structure of Tone Units

The prosodic features discussed above correspond nicely with those used by Crystal (1969) in his analysis of the phonological segmentation of spoken sentences. Crystal's work is based on the transcription of an extensive data base of adult spontaneous speech. The analysis describes the organization of speech in terms of tone units, which were defined by prosodic rather than syntactic cues. A tone unit is characterized by a single peak of pitch prominence or tonicity which usually falls on the last lexical item. The boundary of the unit is marked by a perceivable pitch change and by the presence of junctural features such as a slight pause and/or segmental modifications such as lengthening or aspiration (p. 206).

Analysis of adult conversational speech showed that tone units are frequently coextensive with single clauses, but not consistently enough to define the clause as the grammatical basis of speech segmentation. The percentage of clauses coextensive with one tone unit was 46%, while the percentage of tone units coextensive with a clause was only 28%. The best generalization about the grammar of tone units is that they normally consist of elements of clause structure or constituents (subject,

verb, complement, adverbial, vocative). Eighty percent of tone units contained one or more of these integral elements of clause structure; e.g., subject, adverbial, verb-complement, subject-verb-complement. Clearly, these elements are the basic building blocks of speech units.

An important finding in Crystal's analysis has to do with the length of units as measured by the number of words they contain. This measure (hereafter called unit-length) is a major feature of tone-unit structure. The number of words in specific elements such as subjects or complements determined whether these would be a) subdivided into separate units, b) set off by themselves, or c) combined intonationally with adjoining elements. In general, tone units tended to have roughly equal lengths, with a mean unit-length of five words. This figure is remarkably consistent with estimates of short-term memory capacity which range from five to seven words (Boomer, 1978; Smith, 1978). This agreement between studies of prosody and verbal memory supports the hypothesis that intonation groups words into "packages" of an appropriate size and structure for the limited capacities of short-term memory.

IMPLICATIONS FOR READING

This review has shown that prosodic cues, such as intonation and timing, assist listeners in the immediate stages of sentence processing by segmenting the speech string into perceptual units. These units are built by constituent elements, but purely syntactic considerations (e.g., height in a tree structure) are not sufficient to specify their boundaries. In particular, intonation contours seem to be governed by an interaction between unit-length and meaning.

The information provided by intonation helps the listener to quickly and efficiently perceive sentence organization and process information. For readers who are poor comprehenders, the loss of cues for tone unit boundaries in printed text may be the most serious difference between oral language and reading. As an instructional aid for such readers, the physical segmentation of text may serve to pre-organize sentences into perceptually functional information units. The rules for constructing these segments seem to be as follows:

a) units should be three to seven words in length.
E.g., As a direct result / of their new invention's influence /

b) units should include integral constituents (NP, V, PP, ADV, or S)
E.g., . . . / [Cher]NP [scolded]V [the children]NP /

c) lexical members of units should be linked together by grammatical relationships (noun-adjunct, subject-verb-object, etc.). For example, a noun-adjunct relationship is shown by [The owner] [of the park] /. An example of a subject-verb-object relationship is . . . / [the company] [was given] [an award] /.

We conclude this discussion by comparing the above guidelines with some judgments given by children as to how sentences should be divided into units on the basis of both sound and meaning. Two example sentences are given below from an excellent study by Cioffi (1980) of children's use and perception of intrasentence units. Using an interviewing technique that explained the task without modeling, Cioffi asked children to segment sentences by placing slash marks in a text where it would be permissible to pause, that is, where pausing would not interfere with meaning. The two examples below show the concensus of good readers in the fourth grade as to where it would be permissible to pause.

(24) The Hawaiian Islands / were formed by lava / slowly flowing from the craters / of volcanoes on the ocean floor.

(25) This morning / Kim was sitting in the forward cabin /
and watching the Earth / through the cabin window.

All of the pause locations identified by the children segment the sentences into phrases that are consistent with the three general rules given above. Each phrase, for example, is between three and seven words in length, except for the preposed adverbial phrase in (25), which is discussed below. In addition, each unit shows a sensitivity to constituent boundaries within sentences. For example, each multiword segment is completed by a noun phrase or a prepositional phrase. And finally, each unit expresses a meaningful relationship. However, not all units are major constituents, such as subject noun phrases. In (24) the subject noun phrase The Hawaiian Islands is set off as a pausal unit, but in (25) the short subject, Kim, is joined, in sound and meaning, with the verb phrase. This is similar to a number of examples already discussed in this paper (e.g., prosodic segmentation in sentences 21 and 22), showing that children's judgments are consistent with the findings of adult studies in speech perception and production.

These sentences also show at least three phenomena which are only partially accounted for by the very general guidelines presented above. In sentence (25), a break occurs at the preposed adverbial phrase This morning, even though this phrase is very short. Although this is accounted for by the restriction that words should reflect immediate semantic relations (which would not be the case in a unit like this morning Kim), sentence adverbial units seem to be set off quite frequently in pausal

judgments, especially if they are preposed.[6] Another interesting aspect of the children's pause placement judgments is that the boundary between underlying sentences is marked in both sentences, even when this boundary is inexplicit in the surface structure (. . . by lava / slowly flowing . . .; . . . cabin / and watching . . .). This suggests that the boundaries of deep sentoids may be especially important in the perceptual segmentation of sentences, as proposed by Fodor, Bever, and Garrett (1974).

Finally, a third feature of the segmentation in (25) points to the importance of subtle aspects of meaning which have not been dealt with in this paper. Cioffi's fourth graders judged a pausal break to be permitted before the prepositional phrase in the second clause (through the cabin window), but not in the first clause (in the forward cabin). This seems to reflect their sensitivity to different co-occurrence restrictions on modifiers attached to verbs like sit and watch: the former seems to require a locative phrase of some kind; but watch requires only a direct object. Other modifiers are possible but optional. The children's strategy for grouping words together takes into account this constraint on English usage. Children's intuitions about sentences suggest that the perceptual effects of constituent preposing, clausal boundaries, and co-occurrence restrictions are important areas for further study.

---

[6] As pointed out to us by Bruce Cronnell, it is likely that any optional, non-basic ordering requires a pause and/or some other junctural feature, e.g., more problems / I don't need, or that / I like. Some support for this view is given by Cooper and Sorensen's (1977) discussion of intonation contours associated with preposed constituents.

In summary, prosodic information gives the listener cues to the perceptual segmentation of sentences. The boundaries of meaningful phrases and clauses are acoustically marked by changes in pitch contour, stress level, and temporal intervals between stress beats. These melodic and rhythmic features, which must be imposed on written text by the reader, constitute reading "with expression." This also implies "reading with comprehension," because prosody reflects major syntactic-semantic relationships which underlie the content of sentences. Because of this close relationship between prosody and meaning, the loss of prosodic information in text may be a major source of difficulty for poor readers. But there is another reason that prosody is important for the young reader. In addition to highlighting meaningful relationships, prosody also serves to "package" sentences in a way that permits efficient allocation of memory resources in sentence processing. This is probably crucial for poor readers whose memory capacities may already be taxed by the attentional demands of decoding operations (Perfetti & Hogaboam, 1975).

Reading instruction traditionally recognizes the importance of print-sound relationships on the level of single word decoding. This review of sentence perception and prosody suggests that print-sound correspondences are also important on the level of phrases, clauses, and sentences. A promising instructional technique for helping children compensate for the lack of prosodic cues in text is the use of phrasally segmented text. Phrasal text provides a visual cue (line boundary) for the segmentation of sentences into meaningful phrases. Moreover, it seems likely that teachers can greatly enhance the value of phrasal text

by modelling its sound in oral reading. The teacher's oral reading could emphasize the features of intonation, stress, and segmental lengthening that perceptually isolate each phrase. Through this kind of experience, children may begin to recognize phrasal units more readily, and to develop strategies for organizing text without the support of explicit auditory or visual cues.

References

Anglin, J. M., & Miller, G. A. The role of phrase structure in the recall of meaningful verbal material. Psychonomic Science, 1968, 10, 343-344.

Bach, E. Syntactic theory. New York: Holt, Rinehart and Winston, 1974.

Bever, T. G., Lackner, J. R., & Kirk, R. The underlying structures of sentences are the primary units of immediate speech processing. Perception and Psychophysics, 1969, 5, 225-234.

Boomer, D. S. Hesitation and grammatical encoding. Language and Speech, 1965, 8, 148-158.

Boomer, D. S. The phonemic clause: Speech unit in human communication. In A. W. Siegmen & S. Feldstein (Eds.), Nonverbal behavior and communication. Hillside, NJ: Erlbaum, 1978.

Carroll, J. M., & Tanenhaus, M. K. Functional clauses and sentence segmentation. Journal of Speech and Hearing Research, 1978, 21, 793-808.

Chomsky, N., & Halle, M. The sound pattern of English. New York: Harper & Row, 1968.

Cioffi, G. L. The perception and use of sentence structure by good compreheners and skilled decoders in contextually limited environments. Ph.D. Dissertation, University of Minnesota, 1980.

Clark, H. H., & Clark, E. V. Psychology and language: An introduction to psycholinguistics. New York: Harcourt, Brace, & Jovanovich, 1977.

Cooper, W. E., Paccia, J. M., & Lapointe, S. Hierarchical coding in speech timing. Cognitive Psychology, 1978, 10, 154-177.

Cooper, W. E., & Sorensen, J. M. Fundamental frequency contours and syntactic boundaries. Journal of the Acoustical Society of America, 1977, 62, 683-692.

Coots, J. H., & Snow, D. P. Comprehension skills and text organization ability in reading. Technical Note 2-81/11. Los Alamitos, CA: SWRL Educational Research and Development, 1981.

Crystal, D. Prosodic systems and intonation in English. Cambridge, England: Cambridge University Press, 1969.

Fodor, J. A., & Bever, T. G. The psychological reality of linguistic segments. Journal of Verbal Learning and Verbal Behavior, 1965, 4, 414-420.

Fodor, J. A., Bever, T. G., & Garrett, M. F. The psychology of language: An introduction to psycholinguistics and generative grammar. New York: McGraw-Hill, 1974.

Garrett, M. F., Bever, T. G., & Fodor, J. A. The active use of grammar in speech perception. Perception and Psychophysics, 1966, 1, 30-32.

Goldman-Eisler, F. Pauses, clauses, sentences. Language and Speech, 1972, 15, 103-113.

Golinkoff, R. A comparison of reading comprehension processes in good and poor comprehenders. Reading Research Quarterly, 1975-1976, 11, 623-659.

Graf, R., & Torrey, J. W. Perception of phrase structure in written language. American Psychological Association Convention Proceedings, 1966, 83-84.

Grimes, J. E. The thread of discourse. The Hague, Holland: Mouton, 1975.

Guthrie, J. T., & Tyler, S. J. Psycholinguistic processing in reading and listening among good and poor readers. Journal of Reading Behavior, 1976, 8, 415-426.

Halliday, M. A. K. Notes on transitivity and theme in English. Journal of Linguistics, 1967, 3, 37-81.

Jackendoff, R. S. Semantic interpretation in generative grammar. Cambridge, MA: MIT Press, 1972.

Jarvella, R. J. Syntactic processing of connected speech. Journal of Verbal Learning and Verbal Behavior, 1971, 10, 409-416.

Kintsch, W., & Vipond, D. Reading comprehension and readability in educational practice and psychological theory. In L.-G. Nilsson (Ed.), Perspectives on memory research: Essays in honor of Uppsala University's 500th anniversary. Hillside, NJ: Lawrence Erlbaum, 1979.

Klatt, D. H. Vowel lengthening is syntactically determined in a connected discourse. Journal of Phonetics, 1975, 3, 129-140.

Ladefoged, P., & Broadbent, D. E. Perception of sequence in auditory events. Quarterly Journal of Experimental Psychology, 1960, 13, 162-170.

Larkey, L. S. Prosodic information in spoken English and its role in perception. University of Minnesota, 1979.

Lea, W. A. Isochrony and disjuncture as aids to syntactic and phonological analysis. Paper presented at the 89th meeting of the Acoustical Society of America, Austin, Texas, 1975.

Lieberman, P. Intonation, perception and language. Cambridge, MA: MIT Press, 1967.

Oaken, R., Wiener, M., & Cromer, W. Identification, organization, and reading comprehension for good and poor readers. Journal of Educational Psychology, 1971, 62, 71-78.

Perfetti, C., & Hogaboam, T. Relationship between single word decoding and reading comprehension skill. Journal of Educational Psychology, 1975, 67, 461-469.

Rubin, A. D. A theoretical taxonomy of the differences between oral and written language. Technical Report No. 35. Urbana, IL: Center for the Study of Reading, University of Illinois, 1978.

Smith, F. Understanding reading: A psycholinguistic analysis of reading and learning to read. New York: Holt, Rinehart, and Winston, 1978.

Stewart, C., & Gough, P. Constituent search in immediate memory for sentences. Proceedings of the Midwestern Psychological Association, 1967.